Auto-Calibration Method to Determine Camera Pose for Stereovision-Based Off-Road Vehicle Navigation

Qi WANG[1], Qin ZHANG[2] and Francisco ROVIRA-MÁS[3]

[1] *Department of Agricultural and Biological Engineering, University of Illinois at Urbana-Champaign, 1304 W. Pennsylvania Avenue, Urbana, IL 61801, USA*
[2] *Department of Biological System Engineering, Washington State University, 24106 N. Bunn Road, Prosser, WA 99350, USA*
[3] *Departamento de Mecanizacion y Tecnologia Agraria, Polytechnic University of Valencia, Campus Camino de Vera, 46022 Valencia, Spain*

Stereo cameras have been used as perception sensors for agricultural vehicle navigation for years. One problem impeding their broader application is the difficulty of calibrating the installation poses of a camera using conventional measuring tools, especially when such a system is used in ill-structured agricultural field environments. The research reported in this paper was aimed to develop an automated calibration method for determining the camera installation pose with respect to a vehicle frame. Based on this method, a binocular stereo camera acquired a sequence of field scenery images as the vehicle moved straight forward for a short distance on a relatively flat surface. An image processing algorithm has been developed to detect some static feature points in the ground image and track their three-dimensional (3D) relative motions with respect to the moving vehicle. A plane best fitting to those detected ground features was then used to determine the camera roll and pitch angles, and the tracked motions of those feature points were used to estimate the camera yaw. Field test results validated that the developed auto-calibration method was capable of determining the camera installation pose at a calibration accuracy of $\pm 1°$ over an approximately 10 m of vehicle traveling distance. The calibrated poses could be used to compensate for the navigation errors induced by camera misalignment.

Keywords : agricultural vehicle navigation, camera installation, camera pose auto-calibration, stereovision

INTRODUCTION

Stereo cameras have been used as perception sensors for agricultural vehicle navigation during the past ten years. Their capability of depth and visual perception provides more flexibility than other 3D perception sensors. Although stereovision demands large computational power, significant advances in processor speed have made 3D visual sensing more practical in agricultural applications (Rovira-Más et al., 2009). A common application of stereovision for agricultural vehicle navigation is the determination of vehicle heading direction, which is important for steering the vehicle to follow a desired path by tracking detected landmarks in field operations. Therefore, an accurate measurement of camera pose, a combination of the roll, pitch, and yaw angles of camera installation with respect to the vehicle frame, is critical for obtaining accurate navigation informa-

---

Corresponding author : Qin Zhang, fax : +1-509-786-9321,
e-mail : qinzhang@wsu.edu

tion from detected visual features. In many agricultural vehicle navigation applications, it is common to install the camera at a forward-looking posture with an appropriate pitch angle pointing downwards to the ground. Because of such installation requirements, it is often difficult to accurately determine the camera pose in calibration processes using conventional measuring methods. Furthermore, the camera pose may also be changed during field operations. Because of these obstacles, it is beyond most vehicle drivers' capability to re-calibrate the camera pose in an agricultural field. An automated calibration process is, therefore, essential to make the stereovision navigation practical for farmer use.

Similar problems exist with stereovision navigation in on-road vehicles. An important difference between agricultural vehicle navigation and on-road vehicle navigation is that the latter normally occurs on a well-structured road with clearly marked boundaries. Those boundaries offer an irreplaceable reference for calibrating the installation pose of a stereo camera (Franke et al., 1998; Coulombeau and Laurgeau, 2003; Labayrade et al., 2003; Collado et al., 2006; Marita et al., 2006; Lamprecht et al., 2007). A few researchers have used a plane fitting approach to estimate the relevant position of camera installation to the road surface (Weber and Atkin, 1997; Sappa et al., 2006). However, there is little research being reported on calibrating camera pose for agricultural vehicle applications. The major challenge for agricultural applications is that the ill-structured terrain introduces great uncertainty into the calibration process. For example, in a planting operation, the field in front of a moving tractor does not have easily distinguishable landmarks to indicate a path. The only available referencing points in the field are the irregular texture of crop residue. This research dealt with these challenges to develop an automated calibration method for accurately estimating the installation pose of the stereovision navigation camera on an agricultural vehicle.

MATERIALS AND METHODS

1. System architecture

The experimental vehicle platform used for this research was a John Deere Gator™ Utility Vehicle (Deere & Co., Moline, IL, USA). A binocular stereo camera (STH-MD1/-C, Videre Design, Menlo Park, CA, USA) was mounted on a rigid frame installed on the front of the vehicle. An onboard computer was used to support the stereovision navigation system.

2. Concept of camera pose automatic calibration

2.1 Definition of stereovision navigation coordinate systems

The stereovision navigation system discussed in this paper employed three coordinate systems: the camera coordinate system $X_cY_cZ_c$, the vehicle coordinate system $X_vY_vZ_v$, and the ground coordinate system $X_gY_gZ_g$ (as illustrated in Fig. 1) to convert visual texture features detectable in field imagery into navigation signals. Specifically, those coordinate systems were defined as follows:

The origin of the camera coordinate system was located at the optical focal point of the left lens (Fig. 2), with the Z axis (Z_c) pointing forward and perpendicular to the left imager, the X axis (X_c) pointing to the right and the Y axis (Y_c) pointing downwards, and both were parallel to the image plane.

The origin of the vehicle coordinate system was the vehicle's center of gravity, with the X axis (X_v) pointing to the front, the Y axis (Y_v) pointing to the right, and the Z axis (Z_v) pointing downwards, as illustrated in Fig. 1.

The ground coordinate system was defined at the starting point of a navigational path, with the origin O_g being located at the intersection of the ground surface and the plumb line that went through the origin O_c. The Y axis (Y_g) pointed to the front, the X axis (X_g) pointed to the right and the Z axis (Z_g) pointed upwards, also illustrated in Fig. 1.
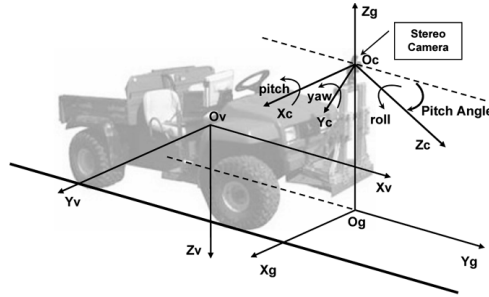
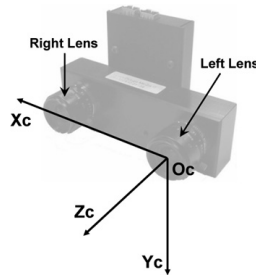**Fig. 1**   Definition illustration of stereovision navigation coordinate systems.

Fig. 2   Definition illustration of camera coordinate system.

2.2   Camera installation pose

In stereovision navigation, the camera *pose* or *attitude*, namely the camera installation orientation, is critical for generating the navigation signals. This is because the visual clues detected by the stereo camera lie in the camera coordinate system, and such clues must be mapped into the vehicle coordinate system or the ground coordinate system before they can be used for navigating the vehicle. The camera pose is often described by the roll, pitch, and yaw angles of the camera with respect to the vehicle.

As mentioned above, a common pose for navigation cameras on agricultural vehicles is looking forward with a pitch angle pointing downwards to the ground. Ideally, the camera should be installed with only the desired downward pitch angle. However, in real installation, it is inevitable for the camera to be installed with small roll and yaw angles with respect to the vehicle coordinate system. Such pose errors will downgrade the navigation accuracy of agricultural vehicles.

2.3   Determination of pose angles

The geometric relationship between the camera and the vehicle coordinate systems is often difficult to accurately determine because of the difficulty in locating the origin O_v. Note, when a vehicle is on a flat surface, the axes of the vehicle and the ground coordinate systems satisfy the following relationships: X_v parallel to Y_g, Z_v parallel to Z_g, and Y_v parallel to X_g as shown in Fig. 1. Therefore, it is possible to directly determine the camera roll, pitch, and yaw angles with respect to the ground coordinate system. Based on the definition of the three coordinate systems, when the camera is installed without roll, pitch, or yaw, there exist three relations:

(1) If roll＝0, the coordinate plane $Y_cO_cZ_c$ is perpendicular to the coordinate plane $X_gO_gY_g$ (the ground surface).

(2) If pitch=0, the coordinate plane $X_cO_cZ_c$ is parallel to the coordinate plane $X_gO_gY_g$.

(3) If yaw＝0, the coordinate plane $Y_cO_cZ_c$ is perpendicular to the coordinate plane $X_gO_gZ_g$.

The three relations lead to a three-step automatic pose calibration process. The basic idea of

this calibration method is to rotate the camera coordinate system from an arbitrary installation pose to an ideal pose without roll, pitch, or yaw. The rotating angles with respect to Z_c, X_c, and Y_c axes indicate the camera roll, pitch, and yaw, respectively.

The determination of the actual camera pose follows the order of roll, pitch, and yaw. In determining the roll angle, the camera coordinate system is rotated around axis Z_c until the coordinate plane $Y_cO_cZ_c$ is perpendicular to the coordinate plane $X_gO_gY_g$, as illustrated in Fig. 3. The opposite of the rotation angle is the roll angle of the camera, β. To find the camera pitch angle, a rotation around axis X_c is followed (Fig. 4). After this rotation, the coordinate plane $X_cO_cZ_c$ of the camera coordinate system is parallel to the coordinate plane $X_gO_gY_g$ in the ground coordinate system, and the opposite of the rotation angle is the pitch angle, α. Finally, a rotation around axis Y_c makes the coordinate plane $Y_cO_cZ_c$ perpendicular to the coordinate plane $X_gO_gZ_g$, which is used to determine the camera yaw angle, ϕ (Fig. 5).

3. Algorithm design

3.1 Overview

Based on the method for determining the camera pose angles introduced above, the developed camera pose auto-calibration process consists of a three-step procedure. In the first two steps, the reference plane used to determine roll and pitch angles is $X_gO_gY_g$, namely the ground surface. Therefore, if the ground surface can be represented in the camera coordinate system, then both the roll and the pitch angles can be calculated. In step 3, the reference plane used to determine the yaw
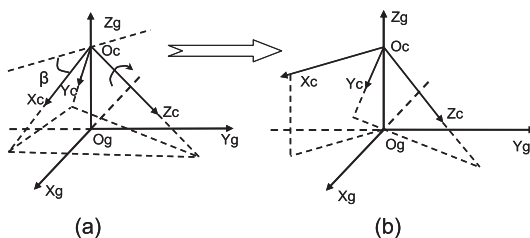

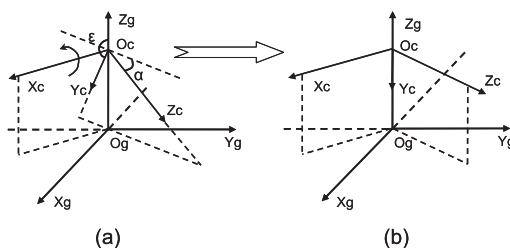Fig. 3   A rotation around axis Z_c to find the camera roll angle.

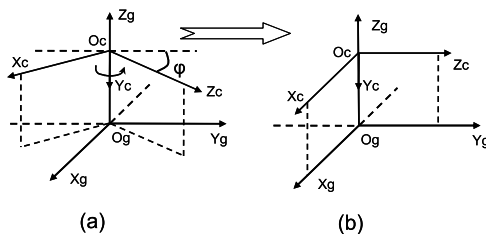Fig. 4   A rotation around axis X_c to find the camera pitch angle.

Fig. 5   A rotation around axis Y_c to find the camera yaw angle.

angle is $X_gO_gZ_g$. When the vehicle platform moves straight forward on a flat surface, its motion direction, which is perpendicular to the coordinate plane $X_gO_gZ_g$, can be used to estimate the alignment error of the camera coordinate system to the ground coordinate system as the indicator of the camera yaw. This auto-calibration method uses vehicle motion direction as a measure to determine the camera yaw angle. Therefore, the camera pose auto-calibration is performed as two tasks: (1) determining a flat ground surface in the camera coordinate system for roll and pitch calibration; and (2) finding the vehicle motion direction in the camera coordinate system for yaw calibration.

In order to comply with the flat surface assumption, this calibration is required to be performed on relatively flat ground. To perform the first task of ground surface determination, it is necessary to know the locations of at least three ground points in the camera coordinate system and then use a plane fitting method to find the equation for ground surface. In the second task of determining the vehicle motion direction, a few static visual landmarks are needed as references for estimating the vehicle motion direction. Unlike the on-road scenario, agricultural fields are generally ill-structured. Therefore, it is difficult to find distinct landmarks, such as traffic lines or road curbs, to serve as the reference landmarks. However, due to the existence of crop residue, farm fields are rich with visual textures, and such textures can be used as landmarks because they can be detected and tracked using a vision system. Plus, those visible textures normally lie on the ground, and can be used to form the ground points for estimating the ground surface. Therefore, ground textures are proper static features to deal with those two tasks.

During the calibration process, the stereo camera continuously shoots field scenery images as the vehicle moves straight forward about 2 m. The developed image processing algorithm identifies and tracks static features in several consecutive image frames to perform the camera pose auto-calibration.

3.2   Ground texture detection

As mentioned above, ground texture is used as a visual clue in finding the static features. The left image in Fig. 6 shows the research platform in a typical farm field covered with soybean residue. When observed from a distance, the residue appears quite homogeneous, which makes it difficult to create a simple rule for converting these residue objects into features indicating the ground surface. Fortunately, at the local detail level, the features of these residues could be readily recognized in terms of their textures. As shown in the right image of Fig. 6, three objects identified by white arrows present distinct corner shapes. Such corner features are commonly seen in residue textures, and therefore are used as static features in this auto-calibration algorithm.

Harris corner detector (Harris and Stephens, 1988) is used to identify residue corner features in this algorithm. As shown in Fig. 7, many corner features can be detected in a field scenery image. This algorithm uses only the detected corner features in the middle lower portion of the image because the distance of the features can be more accurately estimated using the stereo camera.

3.3   Corner feature tracking

In this auto-calibration, a critical step is to track the detected corner features across two consecutive image frames. Corner features are marked as points (pixels) in the tracking process. Because of the large number of pixels in an image (a total of 76,800 pixels for a 320×240 image), plus each pixel having a value between 0 and 255 for its gray scale, it may contain far too many pixels in one image to match one pixel from another image. This algorithm handles this challenge by defining a small window around the point of interest (Ma et al., 2005). Based on this approach, pixel matching becomes window matching, and correspondence among windows across two images is calculated to identify the best matching windows in the tracking process.

Because features move within the field of view in an image as the camera moves during the image acquisition and the image plane is unparallel to the ground surface, the surrounding area of a feature within a fixed range will be deformed between consecutive image frames. Such a
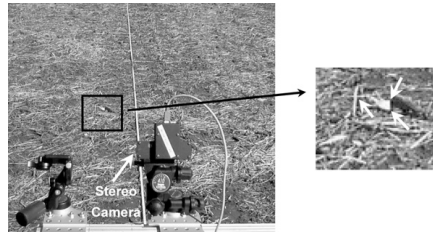
Fig. 6 Soybean residue texture in a typical farm field. Left image shows the vehicle platform with stereo camera in the fields, with the white straight line as a reference line for steering. Right image is a zoom-in view of the rectangle window in left image.
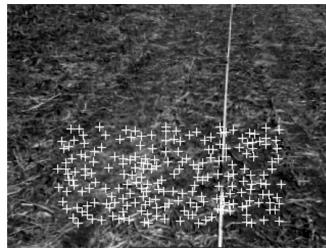

Fig. 7 Corner features (the cross marks) detected in a typical field scenery image frame.

deformation often contributes to window mismatches. This problem is solved by limiting the vehicle speed during calibration to ensure at least four image frames per meter of traveling distance. This speed limitation reduces the deformation between consecutive images so that it is negligible in actual implementations.

In the matching process, the feature points in consecutive image frames are detected separately. An attempt is made to match each feature point in one image with every feature point within a fixed mapping distance of it in the next image (Nister et al., 2006). The longer the mapping distance is, the more computational power is required for matching. Therefore, selecting a proper mapping distance is an important step for efficient feature matching. In this research, a mapping distance of 50 pixels was found adequate for an effective auto-calibration.

To find matching windows in consecutive images, this research used the Normalized Cross-Correlation (NCC) method (Ma et al., 2005) to determine the degree of similarity between windows in interested image frames. The size of the matching window is typically set between 5×5 and 21×21 pixels. A larger window can increase the robustness at the price of higher computational cost. An 11×11 window can offer an acceptable balance between robustness and efficiency in this particular application.

Figure 8 is an example of feature tracking results in a typical field scenery image obtained by using NCC method. While most of the feature pairs show homogenous orientations in the image, indicating the camera motion direction, there are a few pairs of outliers that present obvious different orientations. To obtain an accurate calibration, it is necessary to remove these outliers before further process.

Literature indicated that Random Sample Consensus (RANSAC; Fischler and Bolles, 1980) had proven to be an effective method for eliminating feature matching outliers (Cheng et al., 2005; Nister et al., 2006; Konolige et al., 2008). This auto-calibration algorithm uses a two-stage RANSAC filter to remove the outliners. According to epipolar geometry, for all of the feature points, there is one fundamental matrix that can be used to transform their two-dimensional (2D) positions in the first image to the second image. On the first stage, a RANSAC process is applied
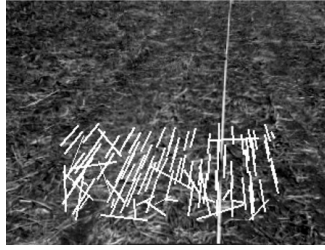
Fig. 8   An example of feature matching results. All of the matched feature pairs between two consecutive image frames are connected by lines.

to delete outliers by fitting a fundamental matrix to the matched feature pairs (Torr, 2002). After removing these outliers, the remaining feature points are transformed from the 2D image space to the 3D camera coordinate system. In the camera coordinate system, all feature points should travel the same distance between two consecutive image frames when the vehicle platform moves straight forward on a relatively flat surface. Therefore, on the second stage, more outliers can be detected by fitting a traveling distance to the matched feature pairs using RANSAC.

In most cases, all outliers can be removed after the application of the two-stage RANSAC process. If necessary, this two-stage process can be repeated. Figure 9 shows that all outliers are removed successfully after applying the two-stage RANSAC filtering procedure.

3.4   Ground surface estimation

The goal of estimating the ground surface is to determine a surface equation in the camera coordinate system. By taking advantage of stereovision capabilities, the ground surface can be readily transformed from a 2D image space to a 3D camera space. Figure 10 depicts the 3D point cloud of the relatively flat ground surface in the camera coordinate system. Theoretically, it requires only three points to determine a plane. Because the ground surface is not perfectly flat, it needs much more than three cloud points to fit the surface with sufficient accuracy. In this auto-calibration algorithm, all the tracked ground features are used to fit the ground surface plane. An orthogonal regression method is used to perform the plane fitting. Figure 11 shows an example of plane fitting results from a typical field scenery image.

3.5   Camera roll angle estimation

The ground surface estimation results in an equation for the ground coordinate plane $X_gO_gY_g$ in the camera coordinate system. Therefore, the angle between coordinate planes $Y_cO_cZ_c$ and $X_gO_gY_g$, named γ, is calculable. The process of roll correction is performed by rotating the camera coordinate system around axis Z_c until γ reaches 90°. The rotating direction can be different due to the property of γ. An obtuse γ angle indicates a negative rotating direction and a sharp γ angle represents a positive rotating direction. An angle θ, defined as $\theta = 90° - \gamma$, is used to determine how far the two coordinate systems are from being perpendicular. When the coordinate plane $Y_cO_cZ_c$ is perpendicular to the coordinate plane $X_gO_gY_g$, θ should be equal to 0. Therefore, θ is a

Fig. 9   Feature tracking results after removing the outliers using the two-stage RANSAC process.
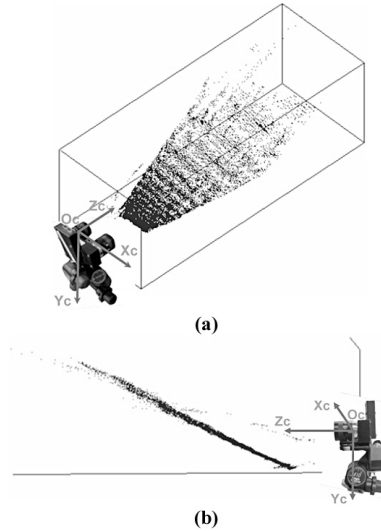
(a)

(b)

Fig. 10 An illustration of 3D point clouds of the ground surface in the camera coordinate system.

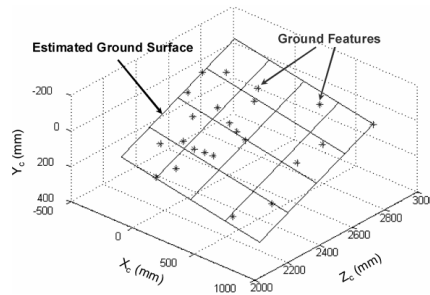**Fig. 11** Example of an estimated ground surface obtained through plane fitting.

control parameter in rotation. Once the absolute value of θ falls below a preset threshold, the roll correction process is completed.

Given the above information, a roll estimation algorithm was developed. Figure 12 shows the implementing flowchart for this algorithm. The co-existence of pitch and yaw made the direct calculation of roll angles complicated. This algorithm applies an indirect approach to solve this problem. It rotates the camera coordinate system around axis Z_c step by step until the absolute value of θ falls below a preset threshold of 0.2°. The estimated camera roll angle is the opposite of the total rotation angle.

3.6 *Camera pitch angle estimation*

The pitch estimation begins under a camera position of no roll error. In reality, it starts with a camera pose with initial roll being adequately compensated. Because a typical camera installation pose on an agricultural vehicle always has a downward-looking pitch angle, the camera pitch estimation method designed in this auto-calibration algorithm is to rotate the camera coordinate system around axis X_c until plane $X_cO_cZ_c$ is parallel to plane $X_gO_gY_g$. The opposite of the rotating angle is treated as the pitch angle of the camera installation pose. Axis Y_c and axis Z_g are the norm vectors of two known planes $X_cO_cZ_c$ and $X_gO_gY_g$ respectively; the angle between the two axes, named ε (Fig. 4), is calculable. Because ε is supplementary to the rotating angle, the actual camera pitch angle is $\alpha = -(180° - \varepsilon)$.
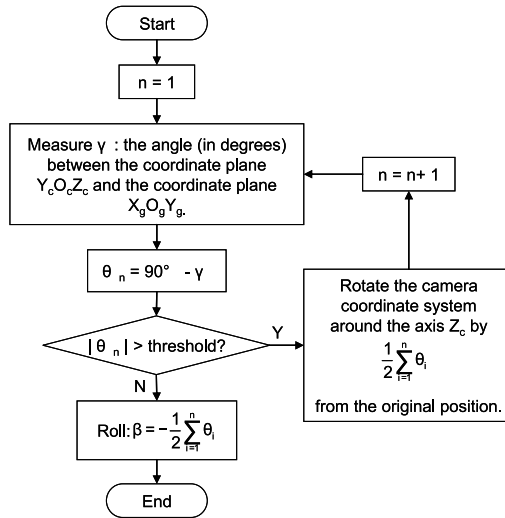
Fig. 12   Implementing algorithm of camera roll estimation.

3.7   Camera yaw angle estimation

The yaw angle estimation is performed through determining vehicle motion direction in the camera coordinate system. By using the previous estimated roll and pitch angles, the camera coordinate system is rotated to a new pose with only the yaw remaining, denoted as $X_{c1}Y_{c1}Z_{c1}$. Figure 13 presents 3D feature motion vectors in the $X_{c1}Y_{c1}Z_{c1}$ space. The variations of the vertical coordinates (Y_{c1}) of the vectors in $X_{c1}Y_{c1}Z_{c1}$ are less than 100 mm, meaning the ground features and their motions are transformed to a relatively level plane. The top view of the system $X_{c1}Y_{c1}Z_{c1}$ (Fig. 14) shows that all motion vectors have almost the same orientation. Their angular deviations from the camera facing direction are caused by the camera yaw angle. Therefore, the average angular deviation of these motion vectors is treated as the camera yaw estimation.

4.   Design of experiments

A series of field experiments were performed to validate the developed method for camera pose auto-calibration. The experimental design focused on four issues: (1) the basic requirements on calibration ground; (2) the camera installation pose; (3) the pose estimation process; and (4) the validation method for pose estimation.

4.1   The basic requirements on calibration ground

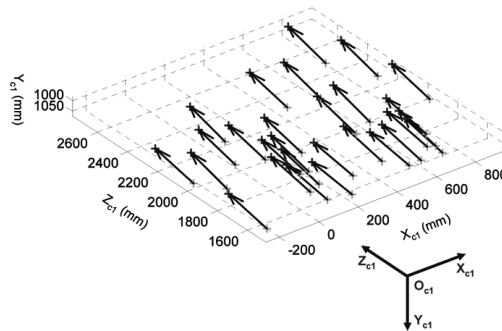The basic requirements on an adequate ground surface for the stereo camera pose auto-

Fig. 13   Feature motion vectors in the 3D space of $X_{c1}Y_{c1}Z_{c1}$. All vectors point to the vehicle heading direction.
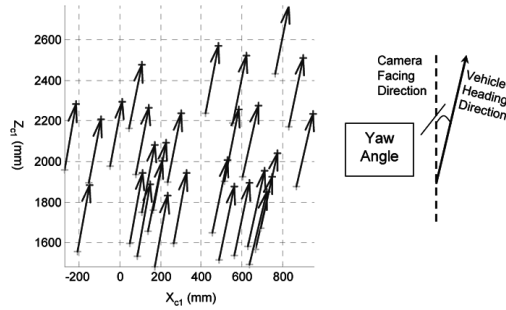
Fig. 14   Top view of feature motion vectors in the 3D space of $X_{c1}Y_{c1}Z_{c1}$.

calibration are: (1) a relatively flat surface about 2 m long without noticeable ditches or bumps; and (2) having a sufficient number of visual corner features on the surface.

These basic calibration ground requirements are not difficult to satisfy. In this research, four types of ground surfaces, including an unploughed soybean field, an unploughed corn field, a grass lawn, and a gravel driveway outside of a garage, were used in validation tests of the developed auto-calibration method. To satisfy the requirement of a relatively flat surface, the selected calibration ground should have no noticeable ditches or bumps. For the second requirement of having sufficient identifiable visual features, a series of preliminary tests showed that lawn texture was too uniform to track; gravel driveway and unploughed soybean/corn fields carried sufficient detectable texture features to support the auto-calibration.

4.2   Camera pose settings

Three camera pose settings were tested to simulate possible situations in actual calibration (Table 1). In setting 1, the camera was installed with a forward-looking pose and a downward pitch angle of 25°, which is a commonly used pose for agricultural vehicle navigation in structured (rowed) crop fields. Settings 2 and 3 were test settings for validating the capability of the auto-calibration method on compensating for pose installation errors. In setting 2, the camera yaw angle was increased by 5° from setting 1, and setting 3 increased another 5° of yaw to setting 2. The reason for changing camera yaw was that errors in yaw have more influence than errors in pitch and roll on vehicle navigation, particularly in finding the heading direction. Because manual installation could not guarantee the camera pose with high accuracy, the roll, pitch, and yaw of each pose fell in an error range of $\pm 2°$.

4.3   Process of pose calibration

During calibration, the vehicle traveled straight forward at a slow and steady speed between 0.5 and 1 ms^{-1} on the selected test ground for about 2 m. Meanwhile, the stereo camera captured scenery images continuously at a rate of four frames per second.

4.4   Validation method for pose estimation

The pose estimation validation consisted of checking whether the calculated pose successfully reflected the relationship between the camera and the vehicle coordinate systems. As mentioned in the calibration algorithm section, static ground features were tracked to estimate the camera yaw. When the camera pose was transformed to a position without roll, pitch, or yaw after a successful

Table 1   Camera pose settings for field tests.

Settings	Camera pose		
	Roll	Pitch	Yaw
1	5°±2°	25°±2°	5°±2°
2	5°±2°	25°±2°	10°±2°
3	5°±2°	25°±2°	15°±2°

calibration, a repetition of the calibration process would result in 0° for the estimated yaw angle. This was the idea implemented in the validation algorithm. A validation ground of 10 m long, extending right beyond the 2 m calibration ground, was used to run such validation tests. Similar to the calibration process, the camera acquired scenery images continuously at a rate of four frames per second in every test run; and the testing vehicle traveled straight forward at a slow and steady speed.

After the scenery images have been acquired, the calibration algorithm was executed to process the images for estimating the camera pose, and then compensated the pose in the navigation algorithm. After the pose was compensated, the validation algorithm was implemented to process the images acquired in the extended validation test to verify the pose compensation result. Figure 15 is the flow chart of the validation algorithm. As shown in this flow chart, the validation algorithm used the first two consecutive image frames taken from the validation test to estimate the compensated pose of the camera by applying the same method as in the calibration process. If the camera pose was properly calibrated, this calculated camera yaw angle should be close to 0° in the validation process. A total of 40 pairs of consecutive image frames, covering about 10 m of traveling distance, were processed in this validation. If the average yaw from those 40 pairs was within a range of ±1°, the camera pose was properly calibrated.

RESULTS AND DISCUSSION

A series of field tests were conducted at the South Farms of the University of Illinois (Urbana, IL, USA) during the spring of 2009. In accordance with the experimental design, a total of 16 field calibration runs, on four types of ground surfaces explained in section 4.1, were conducted. Four test runs with different camera pose installations were conducted on each type of ground. In the first run on each test ground, the camera was installed according to pose setting 1 as defined in section 4.2. The camera poses at the rest runs were on settings 1, 2 and 3, respectively.

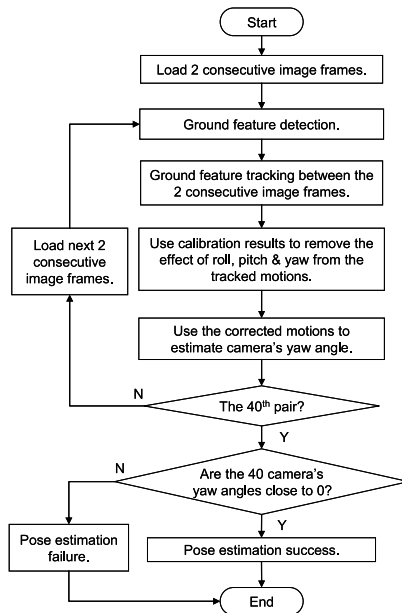Because the calibration process failed for the four runs conducted on lawn surfaces, only 12

Fig. 15 Flow chart of camera pose estimation validation algorithm.

validation runs were performed in this study. The estimated poses obtained from calibration tests on setting 1 showed the absolute values of the estimated roll and yaw angles were very close to 5°, and the estimated pitch angles were between 23° and 24°. At pose settings 2 and 3, it was verified that the yaw angles increased continuously by a value between 4° and 6° from setting1, and matched the actual yaw increase of about 5° between the settings. Figure 16 shows the validation result obtained from a run with camera pose setting 2 in an unploughed soybean field. The result showed that the compensated yaw angles obtained from this validation test were close to 0°. Statistics showed that the average yaw angle was 0.2°, and fell within the preset acceptable error range of $\pm 1°$. Similar to this run, all other validation runs resulted in an average compensated yaw angle within the preset $\pm 1°$ band. Based on the validation results, it can be concluded that the calibration process is able to be automatically implemented on unploughed soybean or corn fields, and on gravel driveways.

Although the field tests had validated the auto-calibration method on three types of farm ground, there were two issues to be examined for better evaluating the developed auto-calibration method: (1) how would the flatness of the calibration field affect the calibration; and (2) how could the ground texture difference affect the calibration?

A series of field validation tests were performed to find a solution for the first issue. In those tests, all selected fields were relatively flat with only some unavoidable roughness typically seen in crop production fields. Because the auto-calibration method was developed under the assumption of a flat surface, such roughness on the ground resulted in small fluctuations on the calculated yaw angle around 0° (Fig. 16). Such fluctuations on yaw angle showed an approximately periodic pattern, which suggests that the effect of roughness could be treated as white noise. A moving average filter could be used to reduce the noise level. Figure 17 shows a noticeable improvement on reducing the fluctuations using a 5-point moving average filter.

To assess the ground flatness on calibration robustness, two tests were conducted in a field with some noticeable bumps. The results showed that, as the ground became rougher, the resulting average yaw angle started to drift from the range of 0.5° to a larger range of 1°. The results shown in Fig. 18 indicate a much larger level of fluctuation on compensated camera yaw angles. This change was attributed to the response of the field scenery images to some large vehicle (camera) rotations induced by the uneven ground surface. This result verified that a good calibration ground should avoid any large bumps/ditches whenever possible. Overall, the calibration results obtained
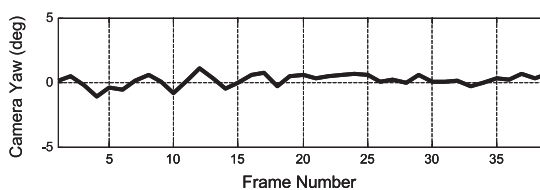

Fig. 16 Compensated camera yaw angles after roll, pitch and yaw correction. The small fluctuation was induced by the rough surface of the calibration ground.
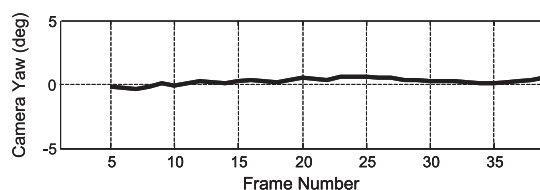

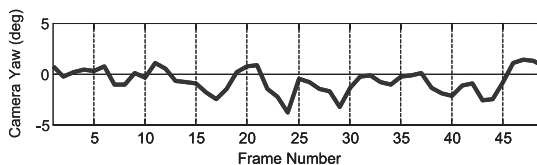**Fig. 17** Smoothed camera yaw angles after applying a 5-point moving average.

Fig. 18 Compensated camera yaw angles after roll, pitch and yaw correction when the calibration was performed on a rough test ground.

from unploughed soybean fields displayed smaller standard deviations in the estimated yaw ($0.5°$ to $0.8°$) than did those from the unploughed corn fields and from the gravel driveways ($1.1°$ to $1.6°$). The possible cause of this difference might be that the unploughed soybean field was less bumpy than the other test fields.

The validation tests also showed that the difference in ground textures could affect the calibration results in different ways. For example, a comparison of the typical texture features on the four types of test grounds used in this research revealed that lawn texture was much more homogeneous (had fewer patterns) than others, which made it very difficult to identify and track ground features in a reliable manner. Consequently, the calibration method failed to perform properly on the lawn surface. It indicated that the calibration runs should avoid lawn covered ground.

CONCLUSIONS

An automatic calibration method was developed to compensate for camera installation errors in a stereovision-based agricultural vehicle navigation system. This method used a set of consecutive field scenery images captured using a vehicle-mounted stereo camera while the vehicle was moving straight forward at a slow speed on relatively flat agricultural grounds, including unploughed soybean and corn fields, lawn ground, and gravel driveways. The tracked ground features were first used to determine the equation of the ground surface in the camera coordinate system for estimating camera roll and pitch angles. Then, the 3D relative motions of these tracked ground features in relation to the camera were used to estimate the camera yaw.

A validation method was also formed to verify the developed calibration algorithm. The validation method could be used to compensate for the camera installation pose with the calibrated roll, pitch and yaw, and then estimate the camera yaw angle after the vehicle traveled straight forward about 10 m. The estimated yaw angle should be close to $0°$ if the camera installation pose calibration results were correct. Field tests showed that the developed calibration method could successfully estimate camera poses as long as there were relatively flat unploughed crop fields or gravel driveways of at least 2 m long.

REFERENCES

Cheng, Y., Maimone, M. W., Matthies, L. 2005. Visual odometry on the mars exploration rovers. Proc. The International Conference on System, Man and Cybernetics, p 903−910.

Collado, J. M., Hilario, C., de la Escalera, A., Armingol, J. M.　2006.　Self-calibration of an on-board stereo-vision system for driver assistance systems. Proc. Intelligent Vehicles Symposium, p 156−162.

Coulombeau, P., Laurgeau, C.　2003.　Vehicle yaw, pitch, roll and 3D lane shape recovery by vision. Proc. Intelligent Vehicles Symposium, p 619−625.

Fischler, M. A., Bolles, R. C.　1980.　Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM **24**: 381−395.

Franke, U., Gavrila, D., Gorzig, S., Lindner, F., Paetzold, F., Wohler, C.　1998.　Autonomous driving goes downtown. IEEE Intelligent Systems & Their Applications **13**: 40−48.

Harris, C., Stephens, M.　1988.　A combined corner and edge detector. Proc. The Fourth Alvey Vision Conference, p 147−151.

Konolige, K., Agrawal, M., Bolles, R. C., Cowan, C., Fischler, M., Gerkey, B.　2008.　Outdoor mapping and navigation using stereo vision. Springer Tracts in Advanced Robotics **39**: 179−190.

Labayrade, R., Aubert, D., Tarel, J.　2003.　Real time obstacle detection in stereovision on non flat road geometry through "v-disparity" representation. Proc. Intelligent Vehicle Symposium, p 646−651.

Lamprecht, B., Rass, S., Fuchs, S., Kyamakya, K.　2007.　Extrinsic camera calibration for an on-board two-camera system without overlapping field of view. Proc. 10th International IEEE Conf. on Intelligent Transportation Systems, p 265−270.

Ma, Y., Soatto, S., Kosecka, J., Sastry, S.　2005.　An invitation to 3-D vision. Springer, New York, NY, USA.

Marita, T., Oniga, F., Nedevschi, S., Graf, T., Schmidt, R.　2006.　Camera calibration method for far range stereovision sensors used in vehicles. Proc. IEEE Intelligent Vehicles Symposium, p 356−363.

Nister, D., Naroditsky, O., Bergen, J.　2006.　Visual odometry for ground vehicle applications. Journal of Field Robotics **23**: 3−20.

Rovira-Más, F., Wang, Q., Zhang, Q.　2009.　Bifocal stereoscopic vision for intelligent vehicles. International Journal of Vehicular Technology. Article ID 123231, 9 p.

Sappa, A. D., Geronimo, D., Dornaika, F., Lopez, A.　2006.　On-board camera extrinsic parameter estimation. Electronics Letters **42**: 745−747.

Torr, P. H. S.　2002.　A structure and motion toolkit in matlab: Interactive adventures in S and M. Microsoft Research Technical Report, Number: MSR-TR-2002-56, pp 40.

Weber, J. W., Atkin, M.　1997.　Further results on the use of binocular vision for highway driving. Proc. Transportation Sensors and Controls: Collision Avoidance, Traffic Management, and ITS, p 52−61.